Table S4. Items and response options relating to risk of reporting biases

Article ID	Tool	Items	Response options
Balshem 2013[1]	AHRQ outcome and analysis reporting bias framework	1. Across all study source documents, what is the risk of ORB/ARB? Compare published report(s) against (1) study protocol (if not retrieved in literature search), (2) trial registry entry/regulatory documents/industry documents, (3) other sources if applicable. 2. If ORB risk unclear: Given the study objectives, duration, and other investigated outcomes, could the study have also likely measured the outcome of interest but not reported it?	**Outcome reporting bias risk positive (ORB risk +)**: If reviewers determine that an outcome X was planned but the results were not reported, or were only partially reported in study documents, then the study is at risk of reporting bias for that outcome ("ORB risk +"). Also, if reviewers determine that an outcome X was not planned but the results were reported, then the study is at risk of reporting bias for that outcome ("ORB risk +"). Also, for studies for which the risk of reporting bias cannot be ruled out, reviewers should ask the question: "Given the study objectives, duration, and other investigated outcomes, could the study have also likely measured the outcome of interest but not reported it?" When the answer is "yes" (e.g., another reported outcome in the study leads the reviewer to believe that outcome X would have been collected), then the study should be rated "ORB risk +" for that outcome. **Outcome reporting bias risk negative (ORB risk -)**: When it is clear to the reviewers that outcome X was planned (e.g. from protocol, regulatory submissions, etc.), complete outcome data are available from at least one study document (published or otherwise), and the outcome was appropriately analyzed as planned, then the study is not at risk for reporting bias for this outcome. Also, for studies for which the risk of reporting bias cannot be ruled out, reviewers should ask the question: "Given the study objectives, duration, and other investigated outcomes, could the study have also

Article ID	Tool	Items	Response options
			likely measured the outcome of interest but not reported it?" If the answer is "no" the study should be rated as "ORB risk–".
			Outcome reporting bias risk unclear (ORB risk unclear): If the reviewers are unable to determine whether an outcome X was planned, but data are reported completely or partially, then the study risk of outcome and analysis reporting bias may be categorized as "unclear". This would also apply to a study that did not report any outcome of review interest across all source documents but was eligible on population, intervention, comparator, and other criteria. Also, for studies for which the risk of reporting bias cannot be ruled out, reviewers should ask the question: "Given the study objectives, duration, and other investigated outcomes, could the study have also likely measured the outcome of interest but not reported it?" If it still remains unclear whether the outcome of interest may have been assessed, the study should be categorized as "ORB risk unclear."
			Analysis reporting bias risk positive (ARB risk +): When reported results are based on a different analysis, effect measure, cut-off, etc. than what was prespecified, then the study is at risk of analysis reporting bias for that outcome ("ARB risk +"). A study is also at risk of analysis reporting ("ARB risk +") because there is no way to know whether the reported analysis was planned or post hoc.
			Analysis reporting bias risk negative (ARB risk -): When it is clear to the reviewers that outcome X was planned (e.g. from protocol, regulatory submissions, etc.),

Article ID	Tool	Items	Response options
			complete outcome data are available from at least one study document (published or otherwise), and the outcome was appropriately analyzed as planned, then the study is not at risk for reporting bias for this outcome **Analysis reporting bias risk unclear (ARB risk unclear)**: If the reviewers are unable to determine whether an outcome X was planned, but data are reported completely or partially, then the study risk of outcome and analysis reporting bias may be categorized as "unclear". This would also apply to a study that did not report any outcome of review interest across all source documents but was eligible on population, intervention, comparator, and other criteria.
Berkman 2013[2]	AHRQ tool for evaluating the risk of reporting bias	1. Are all the following criteria met: ≥10 studies contributing data for an outcome, studies of unequal sizes, no substantial clinical and methodological differences between smaller and larger studies, and quantitative results accompanied with measures of dispersion? 2. If yes, do smaller studies tend to demonstrate more favorable results? (visual assessment) 3. If yes, what is the result of a test for funnel plot asymmetry? 4. If test is positive, would a clinical decision differ for estimates from a fixed effects versus random effect model because the findings from a fixed effect model are closer to the null? 5. If no to the first question, is there an explanation for substantial heterogeneity?	**Suspected risk of reporting bias**: Testing for funnel plot asymmetry demonstrates a substantial likelihood of bias, and/or a qualitative assessment suggests the likelihood of missing studies, analyses, or outcomes data that may alter the conclusions from the reported evidence. **Undetected risk of reporting bias**: All alternative scenarios.

Article ID	Tool	Items	Response options
		6. If no to any of Q1-5, what is the estimated N of studies that are affected by SOR, SAR, nonpublication, or nonaccessibility? 7. If no to any of Q1-5, what is the total sample size of evidence affected by reporting bias (when known)? 8. If no to any of Q1-5, what is the total N of studies in evidence base? 9. If no to any of Q1-5, what is the total N of participants in evidence base? 10. If no to any of Q1-5, what is the consistency of effect estimates across contributing studies? 11. If no to any of Q1-5, what are the study limitations for the evidence base? 12. If no to any of Q1-5, what is the comprehensiveness of study retrieval and identification?	
Downes 2016[3]	AXIS tool (Appraisal tool for Cross-Sectional Studies)	1. Were the results for the analyses described in the methods, presented?	**Yes**: Not stated **No**: Not stated **Do not know/comment**: Not stated
Downs 1998[4]	Downs-Black tool	1. If any of the results of the study were based on "data dredging", was this made clear?	**Yes**: Any analyses that had not been planned at the outset of the study were clearly indicated. Also, no retrospective unplanned subgroup analyses were reported. **No**: Any analyses that had not been planned at the outset of the study were not clearly indicated. **Unable to determine**: Not stated
Guyatt 2011[5-9]	GRADE	1. Study limitations (including selective outcome reporting) 2. Publication bias	**Study limitations domain – No serious limitations, do not downgrade**: Most information is from studies at low risk of bias (i.e. those with low risk of bias for all key

Article ID	Tool	Items	Response options
			criteria, including lack of allocation concealment, lack of blinding, incomplete accounting of patients and outcome events, selective outcome reporting bias, other limitations [stopping early for benefit, use of unvalidated outcome measures, carryover effects in crossover trial, recruitment bias in cluster-randomized trial])
			Study limitations domain – Serious limitations, rate down one level (i.e., from high to moderate quality): Most information is from studies at moderate risk of bias
			Study limitations domain – Very serious limitations, rate down two levels (i.e., from high to low quality or moderate to very low): Most information is from studies at high risk of bias. Selective reporting is present if authors acknowledge prespecified outcomes that they fail to report or report outcomes incompletely such that they cannot be included in a metaanalysis. One should suspect reporting bias if the study report fails to include results for a key outcome that one would expect to see in such a study or if composite outcomes are presented without the individual component outcomes.
			Publication bias domain – Undetected: None of the criteria for "strongly suspected" are met
			Publication bias domain – Strongly suspected: "In general, review authors and guideline developers should consider rating down for likelihood of publication bias when the evidence consists of a number of small studies. The inclination to rate down for publication bias should increase if most of those small studies are industry sponsored or likely to be industry sponsored (or if the investigators share another conflict of interest)...Another

Article ID	Tool	Items	Response options
			criterion for publication bias is the pattern of study results. Suspicion may increase if visual inspection demonstrates an asymmetrical rather than a symmetrical funnel plot or if statistical tests of asymmetry are positive. Although funnel plots may be helpful, review authors and guideline developers should bear in mind that visual assessment of funnel plots is distressingly prone to error. Enhancements of funnel plots may (or may not) help to improve reproducibility and validity associated with their use...Furthermore, systematic review and guideline authors should bear in mind that even if they find convincing evidence of asymmetry, publication bias is not the only explanation. For instance, if smaller studies suffer from greater study limitations, they may yield biased overestimates of effects. Another explanation would be that, because of a more restrictive (and thus responsive) population, or a more careful administration of the intervention, the effect may actually be larger in the small studies...More compelling than any of these theoretical exercises is authors' success in obtaining the results of some unpublished studies and demonstrating that the published and unpublished data show different results. In these circumstances, the possibility of publication bias looms large. The risk of publication bias is probably larger for observational studies than for RCTs, particularly small observational studies and studies conducted on data collected automatically (e.g. in the electronic medical record or in a diabetes registry) or data collected for a previous study. In these instances, it is difficult for the reviewer to know if the observational studies that appear in the literature represent all or a

Article ID	Tool	Items	Response options
			fraction of the studies conducted, and whether the analyses in them represent all or a fraction of those conducted. In these instances, reviewers may consider the risk of publication bias as substantial" [6]. "Guideline panels and authors of systematic reviews should consider the extent to which they are uncertain about the magnitude of the effect due to selective publication of studies and they may downgrade the quality of evidence by one level. Consider: study design (experimental vs. observational); study size (small studies vs. large studies); lag bias (early publication of positive results); search strategy (was it comprehensive?); asymmetry in funnel plot" [8]. "Relevant content: whether publication bias is undetected or suspected; interpretation of funnel plot; comprehensiveness of the search strategies and methods to identify all available evidence; presence of small (often positive) studies with for profit interest...Indicate the reason publication bias is detected (e.g. asymmetrical funnel plot, small studies with positive results, suspected selective availability of data from published, or unpublished studies)" [9].
Hayden 2013[10]	QUIPS (Quality In Prognosis Studies) tool	1. Statistical analysis and reporting (the statistical analysis is appropriate and all primary outcomes are reported). Prompting items include (a) Sufficient presentation of data to assess the adequecy of the analytic strategy; (b) Strategy for model building is appropriate and is based on a conceptual framework or model; (c) The selected statistical model is adequate for the design of the study; (d) There is no selective reporting of results.	**Low risk of bias**: The reported results are unlikely to be spurious or biased related to analysis or reporting **Moderate risk of bias**: The reported results may be spurious or biased related to analysis or reporting **High risk of bias**: The reported results are very likely to be spurious or biased related to analysis or reporting

Article ID	Tool	Items	Response options
Higgins 2008[11-13]	Cochrane risk of bias tool for randomized trials	1. Are reports of the study free of suggestion of selective outcome reporting? (2008 version); Reporting bias due to selective outcome reporting (2011 version)	**Low risk of bias**: Any of the following – The study protocol is available and all of the study's pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way; The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified (convincing text of this nature may be uncommon). **High risk of bias**: Any one of the following – Not all of the study's pre-specified primary outcomes have been reported; One or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not prespecified; One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected adverse effect); One or more outcomes of interest in the review are reported incompletely so that they cannot be entered in a meta-analysis; The study report fails to include results for a key outcome that would be expected to have been reported for such a study. **Unclear risk of bias**: Insufficient information to permit judgement of 'Low risk' or 'High risk'. It is likely that the majority of studies will fall into this category.
Higgins 2016[14 15]	RoB 2.0	1. Are the reported outcome data likely to have been selected, on the basis of the results, from multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain, or from multiple analyses of the data?	**Low risk of bias**: Reported outcome data are unlikely to have been selected, on the basis of the results, from multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain, and reported outcome data are unlikely to have been selected, on the basis of the results, from multiple analyses of the data.

Article ID	Tool	Items	Response options
			High risk of bias: Reported outcome data are likely to have been selected, on the basis of the results, from multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain, or from multiple analyses of the data (or both). **Some concerns**: There is insufficient information available to exclude the possibility that reported outcome data were selected, on the basis of the results, from multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain, or from multiple analyses of the data.
Hoojimans 2014[16]	SYRCLE's RoB tool (SYstematic Review Centre for Laboratory animal Experimentation)	1. Are reports of the study free of selective outcome reporting? Includes two signalling questions: Was the study protocol available and were all of the study's pre-specified primary and secondary outcomes reported in the current manuscript?; Was the study protocol not available, but was it clear that the published report included all expected outcomes (i.e. comparing methods and results section)?	**Low risk of bias**: Not stated, but assume same criteria as Cochrane risk of bias tool for randomized trials [13]. **High risk of bias**: Not all of the study's pre-specified primary outcomes have been reported; One or more primary outcomes have been reported using measurements, analysis methods or data subsets (e.g. subscales) that were not pre-specified in the protocol; One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting has been provided, such as an unexpected adverse effect); The study report fails to include results for a key outcome that would be expected to have been reported for such a study. **Unclear risk of bias**: Not stated, but assume same criteria as Cochrane risk of bias tool for randomized trials [13].
Kim 2013[17]	RoBANS (Risk of Bias Assessment	1. Reporting biases caused by the selective reporting of outcomes	**Low risk of bias**: Any one of the following conditions – The experimental protocol is available, and the pre-

Article ID	Tool	Items	Response options
	Tool for Nonrandomized Studies)		defined primary/secondary outcomes were described as planned; All of the expected outcomes were included in the study descriptions (even in the absence of the experimental protocols). **High risk of bias**: Any one of the following conditions – The pre-defined primary outcomes were not fully reported; The outcomes were not reported in accordance with the previously defined standards; Primary outcomes that were not pre-specified in the study existed (except for outcomes with clear explanations, such as unexpected adverse effects); The existence of incomplete reporting regarding the primary outcome of interest; The absence of reports on important outcomes that would be expected to be reported for studies in related fields. **Unclear risk of bias**: It is uncertain whether the selective outcome reporting resulted in a 'high risk' or a 'low risk' of bias.
Kirkham 2010[18 19]	ORBIT-I (Outcome Reporting Bias In Trials) classification system for benefit outcomes	1. The Outcome Reporting Bias In Trials (ORBIT) study classification system for missing or incomplete outcome reporting in reports of randomised trials	**Low risk of bias**: A “low risk” classification was awarded when it was suspected, but not actually known, that the outcome was either not measured, measured but not analysed, or measured and analysed but either partially reported or not reported for a reason unrelated to the results obtained. Specific examples include: (C) Trial report states that outcome was analysed but insufficient data were presented for the trial to be included in meta-analysis or to be considered to be fully tabulated; (F) Clear that outcome was measured but not necessarily analysed, and judgment says unlikely to have been analysed but not reported because of non-significant

Article ID	Tool	Items	Response options
			results; (H) Not mentioned but clinical judgment says outcome unlikely to have been measured at all. **High risk of bias**: A "high risk" classification was awarded when it was either known or suspected that the results were partially or not reported because the treatment comparison was statistically non-significant (P>0.05). Specific examples include: (A) Trial report states that outcome was analysed but only reports that result was not significant (typically stating P>0.05); (D) Trial report states that outcome was analysed but no results reported; (E) Clear that outcome was measured but not necessarily analysed, and judgment says likely to have been analysed but not reported because of non-significant results; (G) Not mentioned but clinical judgment says outcome likely to have been measured and analysed but not reported on the basis of non-significant results. **No risk of bias**: A "no risk" classification was reserved for cases where it was known that the outcome was not measured, known that it was measured but not analysed, or known that it was measured and analysed but the reason for partial or no reporting was not because the results were statistically non-significant. Specific examples include: (B) Trial report states that outcome was analysed but only reports that result was significant (typically stating P<0.05); (I) Clear that outcome was not measured.
Meader 2014[20 21]	SAQAT (Semi-Automated	Study limitations domain	**Study limitations domain – No serious limitations**: No problem for any source of risk of bias.

Article ID	Tool	Items	Response options
	Quality Assessment Tool)	1. Were data reported consistently for the outcome of interest (i.e. no potential selective reporting)? Publication bias domain 1. Did the authors conduct a comprehensive search? 2. Did the authors search for grey literature? 3. Authors did not apply restrictions to study selection on the basis of language? 4. There was no industry influence on studies included in the review? 5. There was no evidence of funnel plot asymmetry? 6. There was no discrepancy in findings between published and unpublished trials?	**Study limitations domain – Serious limitations**: Selection bias results in serious limitations, or very serious limitations if combined with a problem from any alternative source; two problems from other sources (e.g. detection bias, attrition bias) result in serious limitations. **Study limitations domain – Very serious limitations**: Selection bias results in serious limitations, or very serious limitations if combined with a problem from any alternative source; three problems result in very serious limitations **Publication bias domain – Strongly suspected**: High probability of publication bias. Responses to each item are entered into a Bayesian network to ascertain the probabilities of each GRADE domain. Publication bias is determined by a combination of discrepancy between published and unpublished studies (yes/no), amount of statistical information (high/intermediate/low), industry influence (yes/no) and search integrity (high/low), with the former carrying greatest weight. That is, the probability of publication bias is always considered high when there is a discrepancy between published and unpublished studies (regardless of responses to other items). **Publication bias domain – Undetected**: Low probability of publication bias (as determined by the Bayesian network described above.

Article ID	Tool	Items	Response options
Reid 2015[22]	Selective reporting bias algorithm	1. Protocol available? 2. Trial registration? 3. Outcomes described? 4. Response from contact with study authors? 5. Outcomes match?	**High risk of bias**: Outcomes are described in the protocol or trial registry or by the review authors when contacted, and they do not match the outcomes reported. **Low risk of bias**: Outcomes are described in the protocol or trial registry or by the review authors when contacted, and they do match the outcomes reported. **Unclear risk of bias**: Outcomes are not described in the protocol or trial registry, or a protocol or trial registry are not available and no response is received from review authors when contacted.
Saini 2014[23]	ORBIT-II (Outcome Reporting Bias In Trials) classification system for harm outcomes	1. ORBIT-II classification system	**Low risk of bias**: Specific examples include: (P3) Explicit specific harm measured and compared across treatment groups, although insufficient reporting for meta-analysis or full tabulation; (T1) Clinical judgement says specific harm likely measured but no events, because specific harm not mentioned but all other specific harms fully reported; (T2) Clinical judgement says specific harm likely measured but no events, because there was no description of specific harms; (U) Specific harm outcome not explicitly mentioned, clinical judgment says unlikely measured (no harms mentioned or reported). **High risk of bias**: In the context of harm outcomes, we awarded classifications for "high risk" outcome reporting bias when the specific harm had been measured but the data were presented or suppressed in a way that would mask the harm profile of particular interventions (including providing detail on the seriousness of the harms)—that is, P1, P2, R, and S classifications. Specific examples include: (P1) States outcome analysed but reported only that P>0.05; (P2) States outcome analysed

Article ID	Tool	Items	Response options
			but reported only that P<0.05; (R1) Clear that outcome was measured but no results reported; (R2) Result reported globally across all groups; (R3) Result reported from some groups only; (S1) Clinical judgment says specific harm outcome likely measured and likely compared across treatment groups, but only pooled adverse events reported (could include specific harm outcome); (S2) Clinical judgment says specific harm outcome likely measured and likely compared across treatment groups, but no harms mentioned or reported. **No risk of bias**: Specific examples include: (Q) Clear that explicit specific harm outcome was measured and clear outcome was not compared; (V) Report clearly specifies that data on specific harm of interest was not measured.
Salanti 2014[24 25]	Framework for evaluating the quality of evidence from a network meta-analysis	1. Study limitations (including selective outcome reporting) evaluated in a specific pairwise effect estimated in network meta-analysis: Determine which direct comparisons contribute to estimation of the NMA treatment effect and integrate risk of bias assessments from these into a single judgment. 2. Publication bias evaluated in a specific pairwise effect estimated in network meta-analysis: Non-statistical consideration of likelihood of non-publication of evidence that would inform the pairwise comparison. Plot pairwise estimates on contour-enhanced funnel plot. 3. Study limitations (including selective outcome reporting) evaluated in treatment ranking estimated in network meta-analysis: Integrate risk of bias assessments from each direct comparison to	**Study limitations domain – No serious limitations, do not downgrade**: Use standard GRADE considerations to inform judgment [7]. **Study limitations domain – Serious limitations, rate down one level (i.e., from high to moderate quality)**: Use standard GRADE considerations to inform judgment [7]. **Study limitations domain – Very serious limitations, rate down two levels (i.e., from high to low quality or moderate to very low)**: Use standard GRADE considerations to inform judgment [7]. **Publication bias domain (evaluated in a specific pairwise effect estimated in network meta-analysis) – Undetected**: Use standard GRADE to inform judgment [6].

Article ID	Tool	Items	Response options
		formulate a single overall confidence rating for treatment rankings. 4. Publication bias evaluated in treatment ranking estimated in network meta-analysis: Non-statistical consideration of likelihood of non-publication for each pairwise comparison. If appropriate, plot NMA estimates on a comparison adjusted funnel plot and assess asymmetry.	**Publication bias domain (evaluated in a specific pairwise effect estimated in network meta-analysis) – Strongly suspected**: "Even after a meticulous search for studies, publication bias can occur and usually it tends to lead to overestimation of an active treatment's effect compared with placebo or other reference treatment. Several approaches have been proposed to generate assumptions about the presence of publication bias, including funnel plots, regression methods and selection models, but each has limitations and their appropriateness is often debated. Making judgements about the presence of publication bias in a network meta-analysis is usually difficult. We suggest that for each observed pairwise comparison, judgements about the presence of publication bias are made using standard GRADE. We recommend that the primary considerations are non-statistical (by considering how likely it is that studies may have been performed but not published) and we advocate the use of contour-enhanced funnel plots, which may help in identifying publication bias as a likely explanation of funnel plot asymmetry. Then, judgements about the direct effects can be summarized to infer about the network estimates by taking into account the contributions of each direct piece of evidence" [24]. **Publication bias domain (evaluated in treatment ranking estimated in network meta-analysis) – Undetected**: Use standard GRADE to inform judgment [6]. **Publication bias domain (evaluated in treatment ranking estimated in network meta-analysis) – Strongly suspected**: "Judgments about the potential impact of

Article ID	Tool	Items	Response options
			publication bias in the ranking of the treatments require, as before, consideration of the comprehensiveness of the search for studies and the likelihood that studies may have been conducted and not published. A statistical approach to detecting bias is offered in certain situations by the comparison-adjusted funnel plot for a network of treatments. In such a plot, the vertical axis represents the inverted standard error of the effect sizes as in a standard funnel plot. However, the horizontal axis represents an adjusted effect size, presenting the difference between each observed effect size and the mean effect size for the specific comparison being made. The use of such a plot is informative only when the comparisons can confidently be ordered in a meaningful way; for example, if all comparisons are of active treatment versus placebo, or all are of a new versus an old drug. Examination of any asymmetry in the plot can help to infer about the possible presence of an association between study size and study effect. Asymmetry does not provide evidence of publication bias, however, since associations between effect size and study size can be due to study limitations or genuine heterogeneity of effects" [24].
Sterne 2016[26]	ROBINS-I (Risk Of Bias In Non-randomized Studies of Interventions) tool	1. Is the reported effect estimate likely to be selected, on the basis of the results, from multiple outcome measurements within the outcome domain, multiple analyses of the intervention-outcome relationship, or different subgroups?	**Low risk of bias**: There is clear evidence (usually through examination of a pre-registered protocol or statistical analysis plan) that all reported results correspond to all intended outcomes, analyses and subcohorts. **Moderate risk of bias**: (i) The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent; and (ii) There is no indication of selection of

Article ID	Tool	Items	Response options
			the reported analysis from among multiple analyses; and (iii) There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results. **Serious risk of bias**: (i) Outcomes are defined in different ways in the methods and results sections, or in different publications of the study; or (ii) There is a high risk of selective reporting from among multiple analyses; or (iii) The cohort or subgroup is selected from a larger study for analysis and appears to be reported on the basis of the results. **Critical risk of bias**: (i) There is evidence or strong suspicion of selective reporting of results; and (ii) The unreported results are likely to be substantially different from the reported results. **No information**: There is too little information to make a judgement (for example, if only an abstract is available for the study).
Viswanathan 2012[27]	RTI Item Bank for Assessment of Risk of Bias and Precision for Observational Studies of Interventions or Exposures	1. Are any important primary outcomes missing from the results? 2. Are any important harms or adverse events that may be a consequence of the intervention/exposure missing from the results?	**Yes (for item on primary outcome)**: No specific criteria stated. Only guidance is "Identify all primary outcomes, including timing of measurement, that one would expect to be reported in the study" **No (for item on primary outcome)**: No specific criteria stated. **Cannot determine (for item on primary outcome)**: No specific criteria stated. **Yes (for item on harm outcome)**: No specific criteria stated. Only guidance is “Identify all important harms,

Article ID	Tool	Items	Response options
			including timing of measurement, that one would expect be reported in the study. Drop if not relevant to body of literature." **Partially (for item on harm outcome)**: No specific criteria stated. **No (for item on harm outcome)**: No specific criteria stated. **Assessment of harms not applicable to this study (for item on harm outcome)**: No specific criteria stated.
Viswanathan 2013[28]	RTI Item Bank for Assessing Risk of Bias and Confounding for Observational Studies of Interventions or Exposures	1. Are any important primary outcomes missing from the results? 2. Are any important harms or adverse events that may be a consequence of the intervention/exposure missing from the results?	**Yes, important outcome(s) missing (for item on primary outcome)**: No specific criteria stated. Only guidance is "Identify all primary outcomes that one would expect to be reported in the study, including timing of measurement." **No important outcome (s) missing (for item on primary outcome)**: No specific criteria stated. **Cannot determine (for item on primary outcome)**: No specific criteria stated. **Yes, important outcomes missing (for item on harm outcome)**: No specific criteria stated. Only guidance is "Identify all important harms that one would expect be reported in the study, including timing of measurement. Drop if not relevant to body of literature." **No important outcomes missing (for item on harm outcome)**: No specific criteria stated. **Assessment of harms not applicable to this study (for item on harm outcome)**: No specific criteria stated.

References

1. Balshem H, Stevens A, Ansari M, et al. Finding grey literature evidence and assessing for outcome and analysis reporting biases when comparing medical interventions: AHRQ and the Effective Health Care Program. (Prepared by the Oregon Health and Science University and the University of Ottawa Evidence-based Practice Centers under Contract Nos. 290-2007-10057-I and 290-2007-10059-I.) AHRQ Publication No. 13(14)-EHC096-EF. Rockville, MD: Agency for Healthcare Research and Quality. November 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

2. Berkman ND, Lohr KN, Ansari M, et al. Chapter 15 Appendix A: A Tool for Evaluating the Risk of Reporting Bias (in Chapter 15: Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update). Methods Guide for Comparative Effectiveness Reviews (Prepared by the RTI-UNC Evidence-based Practice Center under Contract No. 290-2007-10056-I). AHRQ Publication No. 13(14)-EHC130-EF. Rockville, MD: Agency for Healthcare Research and Quality. November 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm

3. Downes MJ, Brennan ML, Williams HC, et al. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ open* 2016;6:e011458.

4. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998;52(6):377-84.

5. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336(7650):924-6.

6. Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence—publication bias. *J Clin Epidemiol* 2011;64(12):1277-82.

7. Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). *J Clin Epidemiol* 2011;64(4):407-15.

8. Schünemann H, Brożek J, Guyatt G, et al. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. [Updated October 2013]. Available from http://gdt.guidelinedevelopment.org/app/handbook/handbook.html.

9. Santesso N, Carrasco-Labra A, Langendam M, et al. Improving GRADE evidence tables part 3: detailed guidance for explanatory footnotes supports creating and understanding GRADE certainty in the evidence judgments. *J Clin Epidemiol* 2016

10. Hayden JA, van der Windt DA, Cartwright JL, et al. Assessing bias in studies of prognostic factors. *Ann Intern Med* 2013;158(4):280-6.

11. Higgins JPT, Altman DG, Sterne JAC. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions. Chichester (UK): John Wiley & Sons 2008:187-241.

12. Higgins JPT, Altman DG, Sterne JAC. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from http://handbook.cochrane.org/.

13. Higgins JPT, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.

14. Higgins JPT, Savović J, Page MJ, et al. Revised Cochrane risk of bias tool for randomized trials (RoB 2.0), Version 20 October 2016. Available from http://www.bristol.ac.uk/population-health-sciences/centres/cresyda/barr/riskofbias/rob2-0/ [accessed 19 September 2017].

15. Higgins JPT, Sterne JAC, Savović J, et al. A revised tool for assessing risk of bias in randomized trials. *Cochrane Methods Cochrane Database of Systematic Reviews* 2016;10(Suppl 1):29-31.

16. Hooijmans CR, Rovers MM, de Vries RBM, et al. SYRCLE's risk of bias tool for animal studies. *BMC Med Res Methodol* 2014;14:43.

17. Kim SY, Park JE, Lee YJ, et al. Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *J Clin Epidemiol* 2013;66(4):408-14.

18. Kirkham JJ, Dwan KM, Altman DG, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010;340:c365.

19. Dwan K, Gamble C, Kolamunnage-Dona R, et al. Assessing the potential for outcome reporting bias in a review: a tutorial. *Trials* 2010;11:52.

20. Meader N, King K, Llewellyn A, et al. A checklist designed to aid consistency and reproducibility of GRADE assessments: development and pilot validation. *Systematic reviews* 2014;3(1):82.

21. Stewart GB, Higgins JP, Schunemann H, et al. The use of Bayesian networks to assess the quality of evidence from research synthesis: 1. *PLoS One* 2015;10(3):e0114497.

22. Reid EK, Tejani AM, Huan LN, et al. Managing the incidence of selective reporting bias: a survey of Cochrane review groups. *Systematic reviews* 2015;4:85.

23. Saini P, Loke YK, Gamble C, et al. Selective reporting bias of harm outcomes within studies: findings from a cohort of systematic reviews. *BMJ* 2014;349:g6501.

24. Salanti G, Giovane CD, Chaimani A, et al. Evaluating the quality of evidence from a network meta-analysis. *PLoS One* 2014;9(7):e99682.

25. Higgins JP, Del Giovane C, Chaimani A, et al. Evaluating the quality of evidence from a network meta-analysis. *Value Health* 2014;17(7):A324.

26. Sterne JA, Hernan MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;355:i4919.

27. Viswanathan M, Berkman ND. Development of the RTI item bank on risk of bias and precision of observational studies. *J Clin Epidemiol* 2012;65(2):163-78.

28. Viswanathan M, Berkman ND, Dryden DM, et al. AHRQ Methods for Effective Health Care. Assessing Risk of Bias and Confounding in Observational Studies of Interventions or

Exposures: Further Development of the RTI Item Bank. Rockville (MD): Agency for Healthcare Research and Quality (US) 2013.